

Prosiding

Vol. 8 No. 1 Tahun 2011

# SNTI 2011

26 November 2011



SNTI

Seminar Nasional Teknologi Informasi



Fakultas Teknologi Informasi  
Universitas Tarumanagara  
Jakarta



## Pengembangan Transliterasi Aksara Jawa Dengan Menggunakan Ibus Framework Dan Ekspresi Reguler

Tri Hartanto Noor Hendratwan <sup>1)</sup> Tedy Setiadi <sup>2)</sup> Dewi Soyusiawati <sup>3)</sup>

Teknik Informatika Universitas Ahmad Dahlan Yogyakarta

Jl. Prof. Dr. Soepomo Janturan Yogyakarta

email : di686hz@yahoo.co.id <sup>1)</sup>, tedys@uad.ac.id <sup>2)</sup>, my\_soyus@yahoo.co.id <sup>3)</sup>

### ABSTRACT

*This study discusses the making of font and transliterators to facilitate the use of Javanese script on the computer. Transliterator is made in two forms to use user groups to the level of knowledge of different Javanese script.*

*The first transliterator form of input method Javanese script that can do the transliteration component syllables of the Latin script to Javanese script. Transliterator is intended for users who already know the rules of writing types and Javanese script. The second transliterator form of application that can perform text transliterated into Latin script from Javanese script and vice versa.*

*This study produced three software modules. First, the Carik.ttf Graphite font that can implement the complex behavior of Javanese script on the rendering process. Second, the Carik-IM input method are made by utilizing the Ibus ibus-table. Third, Carik-Pad an transliterator application which utilize regular expressions to describe the text in the process of transliteration.*

### Key words

*Transliterator, Ibus input framework, regular expressions.*

### 1. Pendahuluan

Aksara Jawa adalah aksara yang digunakan dalam sistem penulisan bahasa Jawa. Aksara ini merupakan turunan dari aksara Brahmi kuno yang berasal dari India dan masih memiliki banyak persamaan dengan aksara-aksara modern yang terdapat di Asia Selatan dan Asia Tenggara. Aksara Jawa telah tercantum dalam sistem pengkodean standar Unicode pada versi 5.2 dengan *code point* dari U+A980 sampai U+A9DF[6].

Penggunaan aksara Jawa pada media elektronik terutama komputer sulit dilakukan. Keadaan ini disebabkan oleh dua hal. Pertama, ketiadaan *font* standar aksara Jawa

sebagai representasi aksara Jawa pada komputer. *Font-font* aksara Jawa yang telah banyak beredar seperti, JG Aksara Jawa, Hanacaraka dan Adjisaka masing-masing menempatkan karakter-karakter aksara Jawa pada susunan *code point* yang berbeda dan tidak sesuai dengan standar yang ditetapkan Unicode. Hal ini menyebabkan *font-font* tersebut tidak saling kompatibel. Kedua, ketiadaan *keyboard* atau *keyboard layout* khusus aksara Jawa yang dapat digunakan untuk menginputkan karakter-karakter aksara Jawa pada komputer.

Penelitian ini bertujuan untuk mempermudah penggunaan aksara Jawa pada komputer dengan membuat *font* untuk aksara Jawa yang sesuai dengan standar Unicode serta pedoman penulisan aksara Jawa dan membuat transliterator yang digunakan untuk menginputkan karakter-karakter aksara Jawa pada komputer. Transliterator pertama berupa *input method* yang dapat mengubah komponen suku kata bahasa Jawa dalam aksara Latin menjadi aksara Jawa. Transliterator kedua berupa aplikasi yang dapat merubah teks bahasa Jawa beraksara Latin menjadi aksara Jawa dan sebaliknya.

*Font* dan transliterator pada penelitian ini berdasarkan pada Pedoman penulisan aksara Jawa sesuai dengan kesepakatan bersama Gubernur D.I. Yogyakarta, Jawa Tengah dan Jawa Timur tentang pembinaan, pengembangan dan pelestarian aksara Jawa. Aturan penulisan yang dibahas terbatas mengenai aturan yang berkaitan dengan penulisan aksara, *sandhangan* dan beberapa tanda baca pada aksara Jawa, serta tidak membahas aturan yang berkaitan dengan pembentukan maupun jenis kata dalam bahasa Jawa. Penulisan aksara Jawa juga disesuaikan dengan urutan komponen suku kata aksara Jawa pada standar Unicode versi 5.2.



## Pengembangan Transliterasi Aksara Jawa Dengan Menggunakan Ibus Framework Dan Ekspresi Reguler

Tri Hartanto Noor Hendratwan <sup>1)</sup> Tedy Setiadi <sup>2)</sup> Dewi Soyusiawati <sup>3)</sup>

Teknik Informatika Universitas Ahmad Dahlan Yogyakarta  
Jl. Prof. Dr. Soepomo Janturan Yogyakarta  
email : di686hz@yahoo.co.id <sup>1)</sup>, tedys@uad.ac.id <sup>2)</sup>, my\_soyus@yahoo.co.id <sup>3)</sup>

### ABSTRACT

*This study discusses the making of font and transliterators to facilitate the use of Javanese script on the computer. Transliterator is made in two forms to use user groups to the level of knowledge of different Javanese script.*

*The first transliterator form of input method Javanese script that can do the transliteration component syllables of the Latin script to Javanese script. Transliterator is intended for users who already know the rules of writing types and Javanese script. The second transliterator form of application that can perform text transliterated into Latin script from Javanese script and vice versa.*

*This study produced three software modules. First, the Cariktf Graphite font that can implement the complex behavior of Javanese script on the rendering process. Second, the Carik-IM input method are made by utilizing the Ibus ibus-table. Third, Carik-Pad an transliterator application which utilize regular expressions to describe the text in the process of transliteration.*

### Key words

*Transliterator, Ibus input framework, regular expressions.*

### 1. Pendahuluan

Aksara Jawa adalah aksara yang digunakan dalam sistem penulisan bahasa Jawa. Aksara ini merupakan turunan dari aksara Brahmi kuno yang berasal dari India dan masih memiliki banyak persamaan dengan aksara-aksara modern yang terdapat di Asia Selatan dan Asia Tenggara. Aksara Jawa telah tercantum dalam sistem pengkodean standar Unicode pada versi 5.2 dengan *code point* dari U+A980 sampai U+A9DF[6].

Penggunaan aksara Jawa pada media elektronik terutama komputer sulit dilakukan. Keadaan ini disebabkan oleh dua hal. Pertama, ketiadaan *font* standar aksara Jawa

sebagai representasi aksara Jawa pada komputer. *Font-font* aksara Jawa yang telah banyak beredar seperti, JG Aksara Jawa, Hanacaraka dan Adjisaka masing-masing menempatkan karakter-karakter aksara Jawa pada susunan *code point* yang berbeda dan tidak sesuai dengan standar yang ditetapkan Unicode. Hal ini menyebabkan *font-font* tersebut tidak saling kompatibel. Kedua, ketiadaan *keyboard* atau *keyboard layout* khusus aksara Jawa yang dapat digunakan untuk menginputkan karakter-karakter aksara Jawa pada komputer.

Penelitian ini bertujuan untuk mempermudah penggunaan aksara Jawa pada komputer dengan membuat *font* untuk aksara Jawa yang sesuai dengan standar Unicode serta pedoman penulisan aksara Jawa dan membuat transliterator yang digunakan untuk menginputkan karakter-karakter aksara Jawa pada komputer. Transliterator pertama berupa *input method* yang dapat mengubah komponen suku kata bahasa Jawa dalam aksara Latin menjadi aksara Jawa. Transliterator kedua berupa aplikasi yang dapat merubah teks bahasa Jawa beraksara Latin menjadi aksara Jawa dan sebaliknya.

*Font* dan transliterator pada penelitian ini berdasarkan pada Pedoman penulisan aksara Jawa sesuai dengan kesepakatan bersama Gubernur D.I. Yogyakarta, Jawa Tengah dan Jawa Timur tentang pembinaan, pengembangan dan pelestarian aksara Jawa. Aturan penulisan yang dibahas terbatas mengenai aturan yang berkaitan dengan penulisan aksara, *sandhangan* dan beberapa tanda baca pada aksara Jawa, serta tidak membahas aturan yang berkaitan dengan pembentukan maupun jenis kata dalam bahasa Jawa. Penulisan aksara Jawa juga disesuaikan dengan urutan komponen suku kata aksara Jawa pada standar Unicode versi 5.2.



## 2. Standar Unicode

Penggunaan *font-font* aksara Jawa sebagai representasi aksara Jawa pada komputer akan efektif ketika *font-font* tersebut saling kompatibel, oleh karena itu *font* yang dibuat haruslah mengacu kepada sebuah standar. Standar Unicode adalah aturan *encoding* karakter universal untuk karakter dan teks tertulis. Unicode mendefinisikan langkah konsisten untuk *encoding* teks multi bahasa sehingga memungkinkan pertukaran data teks secara internasional dan menciptakan landasan bagi perangkat lunak global. Unicode menyediakan kapasitas untuk mengencode seluruh karakter yang digunakan dalam bahasa tertulis di seluruh dunia. Jumlah karakter yang bisa diencode mencapai lebih dari satu juta karakter[6].

Aksara Jawa resmi dimasukkan dalam sistem pengkodean standar Unicode pada versi 5.2 dengan *code space* dari U+A980 sampai U+A9DF.

### 2.1. Aksara Jawa

Aksara Jawa merupakan aksara yang bersifat silabik (kesukukataan), dimana masing-masing aksara mewakili sebuah suara. Konsonan pada aksara Jawa memiliki bunyi vokal /a/ atau /o/ sebagai bunyi vokal bawaan[2]. Konsonan aksara Jawa memiliki bentuk alternatif yang digunakan untuk menghubungkan suku kata tertutup konsonan dengan suku kata berikutnya. Bentuk alternatif tersebut disebut sebagai *pasangan*. Konsonan pada aksara Jawa terdiri dari tiga jenis aksara, yaitu;

- Aksara *carakan* yang merupakan aksara pokok dalam sistem penulisan aksara Jawa.
- Aksara *murda* yang hanya dapat dipakai nama, gelar, nama tempat dan nama lembaga.
- Aksara *rekaan* digunakan untuk menuliskan konsonan pada kata-kata bahasa asing yang dipertahankan seperti aslinya.

Tabel 1. Konsonan pada aksara Jawa

Aksara Carakan				
Ha ꦲ ... ꦲ	Na ꦒ ... ꦒ	Ca ꦕ ... ꦕ	Ra ꦫ ... ꦫ	Ka ꦏ ... ꦏ
Da ꦢ ... ꦢ	Ta ꦠ ... ꦠ	Sa ꦱ ... ꦱ	Wa ꦮ ... ꦮ	La ꦭ ... ꦭ
Pa ꦥ ... ꦥ	Dha ꦢꦲ ... ꦢꦲ	Ja ꦗ ... ꦗ	Ya ꦪ ... ꦪ	Nya ꦤꦺ ... ꦤꦺ
Ma ꦩ ... ꦩ	Ga ꦒ ... ꦒ	Ba ꦧ ... ꦧ	Tha ꦠꦲ ... ꦠꦲ	Nga ꦒꦤ ... ꦒꦤ

ꦲ ... ꦲ	ꦒ ... ꦒ	ꦕ ... ꦕ	ꦮ ... ꦮ	ꦏ ... ꦏ
Aksara Murda				
Na ꦒ ... ꦒ	Ka ꦏ ... ꦏ	Ta ꦠ ... ꦠ	Sa ꦱ ... ꦱ	Pa ꦥ ... ꦥ
Ga ꦒ ... ꦒ	Ba ꦧ ... ꦧ			
Aksara Rekaan				
Kha ꦏꦲ ... ꦏꦲ	Dza ꦢꦴ ... ꦢꦴ	Fa/Va ꦱꦠ ... ꦱꦠ	Za ꦴꦴ ... ꦴꦴ	Gha ꦒꦲ ... ꦒꦲ

Aksara Jawa juga mengenal aksara *swara* yang digunakan untuk menuliskan suku kata vokal terutama yang berasal dari bahasa asing dan atau nama untuk mempertegas pelafalannya. Angka dengan sistem desimal yang digunakan untuk menyatakan lambang bilangan atau nomor serta beberapa tanda baca yang memiliki fungsi beragam dan bahkan tidak ditemui dalam aksara Latin.

Tabel 2. Vokal independen pada aksara Jawa

Aksara Swara				
A ꦲ	I ꦲ	U ꦲ	E ꦲ	O ꦲ

Tabel 3. Angka pada aksara Jawa

Angka Jawa									
0 ꦲ	1 ꦲ	2 ꦲ	3 ꦲ	4 ꦲ	5 ꦲ	6 ꦲ	7 ꦲ	8 ꦲ	9 ꦲ

Tabel 4. Tanda baca pada aksara Jawa

Tanda Baca Jawa		
Pada Adeg-Adeg ꦲ	Pada Guru ꦲ	Pada Pancak ꦲ
Pada Lingsa ꦲ	Pada Lungsi ꦲ	Pada Pangkat ꦲ
Pada Lahir ꦲ	Pada Madya ꦲ	Pada Andhap ꦲ
Purwapada ꦲ	Wasanapada ꦲ	Madhyapada ꦲ

Konsonan pada aksara Jawa dapat diubah bunyi bawanya dengan menggunakan *sandhangan*.



Berdasarkan fungsinya *sandhangan* dibagi menjadi tiga, yaitu:

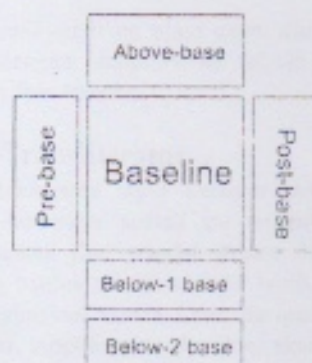
- Sandhangan swara* merupakan tanda diakritik yang digunakan untuk mengubah bunyi vokal pada konsonan aksara Jawa.
- Sandhangan punyigeging wanda* digunakan untuk menghilangkan bunyi vokal bawaan pada konsonan aksara Jawa.
- Panjinan* digunakan untuk menyisipkan bunyi konsonan medial pada konsonan yang masih memiliki bunyi vokal bawaan.

Tabel 5. Sandhangan

Sandhangan Swara			
Wulu (u)	Pepet (e)	Suku (i)	
Taling (e dan i)	Taling Tarung (a)		
Sandhangan Punyigeging Wanda			
Pangkon	Wingnyan	Layar	Cecak
Panjinan			
Cakra (-n-)	Keret (-re)	Pangkal (-v-)	

## 2.2. Perilaku Kompleks Aksara Jawa

Sistem penulisan aksara Jawa lebih kompleks jika dibandingkan sistem penulisan aksara Latin. Kompleksitas penulisan aksara Jawa terlihat pada spesifikasi dan aturan penulisan aksara Jawa baik pada standar Unicode maupun pada pedoman penulisan aksara Jawa.



Gambar 1. Posisi penulisan karakter aksara Jawa

Kompleksitas pertama berkaitan dengan pengaturan posisi dalam penulisan aksara Jawa yang secara visual dapat dibagi dalam beberapa area yaitu:

- Baseline area*, merupakan basis utama penulisan aksara Jawa. Konsonan, vokal independen, angka serta tanda baca di tulis pada area ini
- Pre-base area* (bagian kiri) dan *Post-base area* (bagian kanan), area ini digunakan untuk menuliskan beberapa *sandhangan* seperti *taling* dan *cakra* untuk kasus tertentu di sebelah kiri serta *tarung*, *pangkon* dan beberapa bentuk *pasangan* konsonan disebelah kanan.
- Above-base area*, area ini digunakan untuk menuliskan beberapa *sandhangan* seperti *wulu*, *pepet*, *cecak*, *layar*, dan *cecak telu*.
- Below-1 base area*, area ini digunakan untuk menuliskan bentuk *pasangan* konsonan, *panjinan* dan *sandhangan suku*.
- Below-2 base area*, area ini digunakan untuk menuliskan *panjinan* dan *sandhangan suku* yang mengikuti pasangan konsonan yang terletak pada *below-1 base*.

Kompleksitas penulisan aksara Jawa juga dapat diamati pada aturan penulisan sebagai berikut:

- Contextual shaping*, perubahan bentuk karakter berdasarkan konteks. Hal ini terjadi dalam beberapa kasus. Pertama, pada perubahan aksara *murda* menjadi aksara *carakan* ketika diikuti *pangkon* karena aksara *murda* tidak boleh *sigeg* (digunakan sebagai penutup suku kata). Kedua, penentuan *pasangan* konsonan, *cecak*, *layar*, *wingnyan*, *pa cerek* dan *nga lelet*. Ketiga, perubahan bentuk pada sebagian *pasangan* ketika diikuti oleh *suku*, *cakra*, *keret* atau *pengkal*. Keempat, perubahan bentuk pada *sandhangan suku*, *cakra*, *keret* dan *pengkal* tergantung karakter sebelum dan sesudah *sandhangan* tersebut.
- Reordering*, pengurutan ulang dilakukan ketika ada perbedaan antara urutan secara *logical* berdasarkan standar Unicode dan urutan *visual* berdasarkan aturan penulisan aksara Jawa. Hal ini terjadi pada penulisan *taling* yang secara *logical* berada di belakang aksara pokok tetapi secara *visual* ditampilkan di depan aksara pokok dan pada penulisan *cakra* pada kasus tertentu yang secara *visual* ditampilkan di depan aksara pokok.



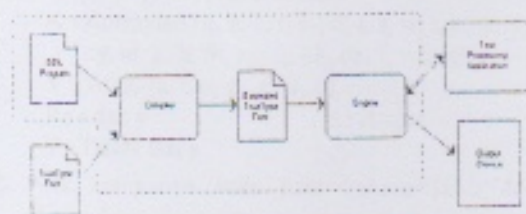
- c. *Stacking diacritics*, pengaturan tanda diakritik yang bertumpuk antara tanda penanda aksara *rekaan* (cecak telu), diakritik *sandhangan swara* (wulu dan pepet) dan diakritik *sandhangan panjigeging wanda* (cecak dan layar).

### 2.3. Graphite

Aksara Jawa memiliki beberapa perilaku kompleks dalam sistem penulisannya oleh karena itu dibutuhkan *smart font* yang dapat mewujudkan perilaku tersebut. *Smart font* adalah font dengan data terlampir yang berisi instruksi untuk menggambarkan bagaimana *glyph-glyph* dipilih dan ditampilkan. Hal ini berbeda dengan font biasa yang memetakan secara langsung antara karakter dan *glyph* dan menampilkan secara berdampingan dalam satu garis lurus[1].

Graphite adalah paket yang dapat digunakan untuk membuat *smart font* yang memiliki kemampuan menangani sistem penulisan dengan berbagai perilaku kompleks, seperti: *contextual shaping*, *ligatures*, *reordering*, *bidirectionality*, *stacking diacritics* dan *complex positioning*[1] [4].

Sistem Graphite terdiri dari bahasa pemrograman *rule-based Graphite Description Language (GDL)* yang digunakan untuk mendeskripsikan perilaku dari sebuah sistem penulisan, *compiler* untuk GDL dan mesin *rendering* yang dapat berfungsi sebagai aplikasi pemrosesan teks[1].



Gambar 2. Arsitektur sistem Graphite[1]

TrueTypeFont biasa dapat diubah menjadi *smart font* dengan mengcompile sebuah program GDL ke dalam font tersebut[1].

### 2.4. Transliterator

Mekanisme input karakter-karakter aksara Jawa pada komputer adalah isu terpenting kedua setelah pembuatan font standar aksara Jawa. Bahasa Jawa dalam bentuk tulisan saat ini lebih sering diwujudkan menggunakan aksara Latin daripada aksara Jawa itu sendiri. Input karakter-karakter aksara Latin juga lebih mudah dibandingkan input karakter-karakter aksara Jawa. Kemudahan ini dapat dimanfaatkan untuk membuat mekanisme input aksara Jawa dalam bentuk transliterator Latin-Jawa.

Transliterator merupakan bentuk kata benda dari kata *transliterate* dalam bahasa Inggris yang berarti *write or print (a letter or word) using the closest corresponding letters of a different alphabet or language* [11]. Transliterasi dalam Bahasa Indonesia berarti penyalinan dengan penggantian huruf dari abjad satu ke abjad yang lain [7].

Transliterasi dapat dilakukan melalui tiga tahap, yaitu tahap penguraian (*parsing*), tahap konversi dan tahap penggabungan. Proses penguraian dilakukan untuk memecah teks yang akan ditransliterasi menjadi kumpulan komponen suku kata bahasa Jawa. Kumpulan komponen suku kata tersebut kemudian dicari padanannya dan dikonversi menjadi aksara tujuan. Kumpulan komponen yang sudah dikonversi kemudian digabungkan kembali sehingga didapat teks hasil transliterasi[8].

Konversi aksara Jawa didasarkan kepada konversi komponen suku kata. Hal ini dilakukan karena adanya perbedaan sifat antara aksara Latin yang bersifat alfabetik dan aksara Jawa yang bersifat silabik. Sedangkan konversi angka Jawa dipetakan 1:1 karena angka pada aksara Jawa menggunakan sistem desimal seperti halnya angka pada aksara Latin. Berdasarkan uraian tersebut maka diperoleh tabel padanan pola komponen suku kata bahasa Jawa sebagai berikut.

Tabel 6. Padanan pola komponen suku kata bahasa Jawa

Aksara Latin	Aksara Jawa
Konsonan	Sigegan (konsonan+pangkon)
Konsonan + a	Konsonan
Konsonan+vokal selain a	Konsonan+sandhangan swara
Konsonan + r + a	Konsonan + cakra
Konsonan + r + vokal selain a	Konsonan + cakra + sandhangan swara
Konsonan+ y + a	Konsonan + pengkal
Konsonan + y + vokal selain a	Konsonan + pengkal + sandhangan swara
A	Ha atau aksara swara a
Vokal selain a	Ha+sandhangan swara / aksara swara
Angka	Angka

Tanda baca pada aksara Jawa berbeda dengan tanda baca pada aksara Latin. Tanda baca pada aksara Jawa dapat membentuk rangkaian tanda baca sehingga menjadi tanda baca lain dengan fungsi yang berbeda dari tanda baca penyusunnya. Sebagian tanda baca pada aksara Jawa memiliki kemiripan fungsi dengan tanda baca pada aksara Latin akan tetapi



sebagian tanda baca yang lain pada aksara Jawa bahkan memiliki fungsi yang tidak ditemukan padanannya pada aksara Latin. Berdasarkan hal tersebut, tabel konversi tanda baca dibuat berdasarkan pada kemiripan fungsi dari tanda baca pada kedua aksara. Tanda baca yang tidak mempunyai padanan fungsi pada aksara Latin dapat diinputkan menggunakan string yang ditentukan. Berikut tabel konversi tanda baca aksara Latin-Jawa yang digunakan sebagai acuan pembuatan metode input aksara Jawa.

Tabel 7. Konversi tanda baca aksara Latin-Jawa

Aksara Latin	Aksara Jawa
,	<i>pada lingsa</i>
..	<i>pada lingsi</i>
: atau ' atau "	<i>pada pangkut</i>
~ \	<i>pada adeg-adeg</i>
~ /	<i>pada windu</i>
~ _	<i>pada undhaw</i>
~ -	<i>pada madyo</i>
~ +	<i>pada luhur</i>

Berdasarkan tabel tersebut maka dapat disusun pola komponen suku kata bahasa Jawa sebagai berikut:

- a. Pola komponen suku kata bahasa Jawa dalam aksara Latin

$\{K\}\{P\}V\}\{A\}\{T\}$ , dimana

K : Konsonan (h, n, c, r, k, d, t, s, w, i, p, dh, j, y, ny, m, g, b, th, ng, q, kh, dz, f, v, z, gh)

V : Vokal (a, i, u, e, é / ê, o)

P : r dan y

A : Angka

T : Tanda Baca

- b. Pola komponen suku kata bahasa Jawa dalam aksara Jawa

$\{K\}\{P\}V\}\{S\}\{B\}\{A\}\{T\}$ , dimana

K : Konsonan S : Pangkon

A : Angka

P : Panjang

B : Aksara swara T : Tanda baca

V : Sandhangan swara

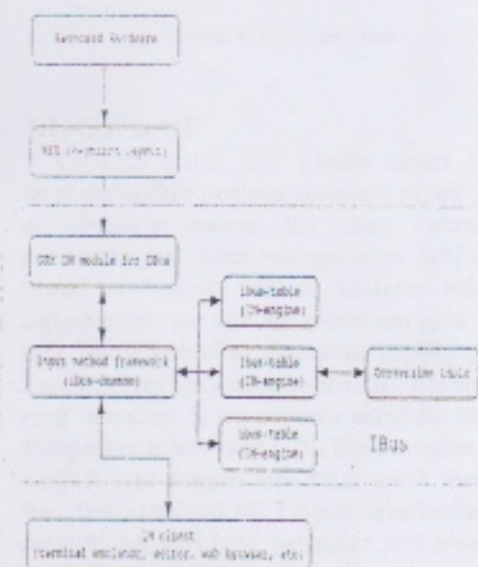
## 2.5. IBus Input Framework

Mekanisme input aksara Jawa dalam bentuk transliterasi Latin-Jawa dapat diwujudkan dengan memanfaatkan IBus *Input Framework*. IBus dipilih karena merupakan *input method framework* yang telah diterima dan digunakan pada FreeBSD dan distro-distro linux besar seperti Debian, Fedora dan Ubuntu. IBus menjadi *input method framework* bawaan sejak Fedora 11 dan Ubuntu 9.10[9].

IBus (*Intelligent input bus*) adalah sebuah *input method framework open source* untuk Linux/UnixOS.

IBus menyediakan tiga macam layanan yaitu *input method engine* (IME), *Panel user interface* (contoh, *language bar*) serta layanan konfigurasi yang menangani konfigurasi untuk IBus dan layanan lainnya seperti IME. IBus memanfaatkan D-Bus untuk mengkomunikasikan ibus-daemon, layanan dan *IM client* seperti, terminal emulator, *text editor* dan *web browser*. ibus-daemon mengatur semua *client* dan layanan dengan menerima registrasi dari layanan kemudian mengirim pesan melalui D-Bus ke *IM client* dan layanan yang sesuai. IBus mendukung beberapa *input method engine* (IME), diantaranya ibus-ant, ibus-chewing, ibus-hangul, ibus-m17n, ibus-pinyin, ibus-unkey dan ibus-table[9].

IBus-table merupakan IME *framework* untuk metode input berbasis tabel. Metode input berbasis tabel yang dimaksud adalah metode input yang hanya mengandalkan tabel konversi untuk menghasilkan huruf sesuai inputan. Metode seperti ini cocok digunakan untuk melakukan transliterasi dari aksara yang bersifat alfabetik menjadi aksara yang bersifat silabik, seperti transliterasi Latin-Jawa.



Gambar 3. Arsitektur IBus

## 2.6. Ekspresi Reguler

Transliterasi dilakukan dengan mengurai teks yang akan ditransliterasi kemudian mengkonversi hasil penguraian tersebut ke aksara tujuan transliterasi. Penguraian (*parsing*) teks asal dapat dilakukan dengan memanfaatkan ekspresi reguler.

Ekspresi reguler dalam teori bahasa formal merupakan notasi yang digunakan untuk mendeskripsikan bahasa reguler. Notasi ekspresi



reguler terdiri dari string simbol alfabet  $\Sigma$ , tanda kurung serta operator  $+$ ,  $.$ , dan  $*$ . Operator  $+$  digunakan untuk melambangkan gabungan, operator  $.$  untuk konkatenasi dan operator  $*$  untuk *star closure*[5].

Definisi formal ekspresi reguler adalah sebagai berikut.

Jika  $\Sigma$  adalah alfabet maka.

- $\emptyset$ ,  $\lambda$  dan  $a \in \Sigma$  adalah ekspresi reguler primitif (sederhana).
- jika  $r_1$  dan  $r_2$  adalah ekspresi reguler, maka demikian juga dengan  $r_1 + r_2$ ,  $r_1 . r_2$ ,  $r_1^*$  dan  $(r_1)$ .
- Sebuah string dapat dianggap sebagai ekspresi reguler jika dan hanya jika dapat diturunkan dari ekspresi reguler primitif dengan menerapkan aturan-aturan pada point b[5].

## 2.7. Java Regex

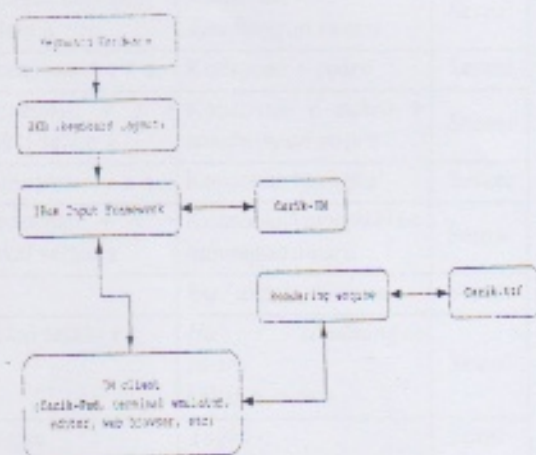
Ekspresi reguler yang digunakan pada proses penguraian bukan hanya sekedar notasi yang digunakan untuk mendeskripsikan bahasa reguler semata. Ekspresi reguler saat ini lebih mirip seperti bahasa pemrograman mini yang digunakan untuk mendeskripsikan dan mengurai teks. Ekspresi reguler juga dapat digunakan untuk melakukan operasi *add*, *remove*, *search*, *replace*, *split* pada setiap jenis teks dan data. [3]. Kemampuan tersebut membuat pemrosesan teks bisa menjadi lebih efisien ketika menggunakan ekspresi reguler.

Ekspresi reguler (*regular expression / regex*) pada bahasa pemrograman Java terdapat dalam paket `java.util.regex`. Pada paket ini terdapat dua kelas yang dapat digunakan untuk mencocokkan urutan karakter dengan sebuah pola yang ditentukan menggunakan ekspresi reguler. Dua kelas tersebut adalah `Pattern` dan `Matcher`. Kelas `Pattern` merupakan representasi terkompilasi dari sebuah ekspresi reguler sedangkan kelas `Matcher` merupakan sebuah mesin yang dapat digunakan untuk melakukan pencocokan urutan karakter dengan menginterpretasi pola yang dihasilkan kelas `Pattern`[10].

## 3. Hasil Penelitian

Penelitian ini menghasilkan tiga *software* yaitu, *font* Graphite Carik.ttf, *input method* Carik-IM dan aplikasi transliterator Carik-Pad. Hubungan ketiga *software* tersebut dapat diketahui melalui proses input karakter aksara Jawa menggunakan IBus. Input melalui *keyboard* menghasilkan *key code* tertentu sesuai dengan tombol yang ditekan. *Key code* tersebut kemudian diproses berdasarkan *keyboard layout* dan diubah menjadi *key symbol*. *Key*

*symbol* ini kemudian diproses lagi menjadi *key symbol* lain berdasarkan *input context* pada IM *client* dengan menggunakan IM *engine* yang aktif. Pada IM *client* *key symbol* yang dihasilkan kemudian diwujudkan secara visual menggunakan *font* yang tersedia oleh *rendering engine*.



Gambar 4. Arsitektur sistem

## 3.1. Carik.ttf

Carik.ttf adalah *font* standar aksara Jawa yang dapat melakukan perilaku kompleks aksara Jawa. *Font* ini dihasilkan melalui dua tahap. Pertama, proses pembuatan *font* dasar menggunakan *fontForge*. Pada tahap ini bentuk seluruh karakter aksara Jawa digambar satu per satu dan diletakkan pada *code point* yang sesuai kemudian digenerate menjadi *font* dasar. Kedua, proses penyisipan kode GDL. Aturan penulisan yang tercakup dalam perilaku kompleks aksara Jawa diwujudkan dalam kode GDL dan kemudian *discompile* menjadi satu dengan *font* dasar untuk menghasilkan *font* Graphite Carik.ttf. Tingkat keberhasilan Carik.ttf dapat dilihat dari hasil pengujian *font* tersebut untuk melakukan perilaku kompleks aksara Jawa.

Tabel 8. Pengujian Carik.ttf

No	Perilaku kompleks	Ket.
Contextual shaping		
1	Aksara <i>marda</i> berubah menjadi aksara <i>carakan</i> ketika diikuti <i>pangon</i>	Berhasil dilakukan
2	<i>Pangkon</i> dan konsonan aksara Jawa yang mengikutinya berubah menjadi bentuk <i>pasangan</i>	Berhasil dilakukan
3	<i>Sigegan ra</i> , <i>nga</i> dan <i>ha</i> berubah	Berhasil



	menjadi <i>layar</i> , <i>cecak</i> dan <i>wingyan</i> .	Dilakukan
4	Aksara <i>ra+pepet</i> dan <i>la+pepet</i> berubah menjadi <i>pa cerek</i> dan <i>nga lelet</i> .	Berhasil dilakukan
5	Perubahan bentuk pada rebanian <i>pusangan</i> ketika diikuti oleh <i>suku</i> , <i>cakra</i> , <i>keret</i> atau <i>pengkal</i>	Berhasil dilakukan
6	Perubahan bentuk pada <i>suku</i> , <i>cakra</i> , <i>keret</i> serta <i>pengkal</i> sesuai karakter yang terdapat pada sebelum dan sesudah <i>sandhangan</i> tersebut	Berhasil dilakukan
<i>Reordering</i>		
7	Penulisan <i>taling</i> yang secara <i>logical</i> berada di belakang aksara pokok tetapi secara visual ditampilkan berada di depan aksara pokok	Berhasil dilakukan
<i>Stacking diacritics</i>		
8	Pengaturan letak tanda diakritik yang bertumpuk antara tanda <i>cecak</i> <i>telu</i> , <i>wulu</i> , <i>pepet</i> , <i>cecak</i> dan <i>layar</i>	Berhasil dilakukan

### 3.2. Carik-IM

Carik-IM merupakan *input method* yang dapat melakukan transliterasi komponen suku kata bahasa Jawa dalam aksara Latin menjadi aksara Jawa. Carik-IM dibuat dengan memanfaatkan Ibus-table yang merupakan bagian dari IBus *input framework*. Pembuatan Carik-IM dilakukan dengan membuat *file* yang mendeskripsikan *input method* kemudian mengkompilasi *file* tersebut. Pada *file* dekripsi tersebut disertakan pula tabel konversi komponen suku kata bahasa Jawa yang akan digunakan oleh ibus untuk melakukan proses *parsing* dan konversi urutan karakter yang diinputkan. Carik-IM dapat digunakan untuk input aksara Jawa menggunakan aksara Latin sesuai dengan tabel konversi. Hal ini dapat dilihat pada hasil pengujian Carik-IM.

Tabel 9. Pengujian Carik-IM

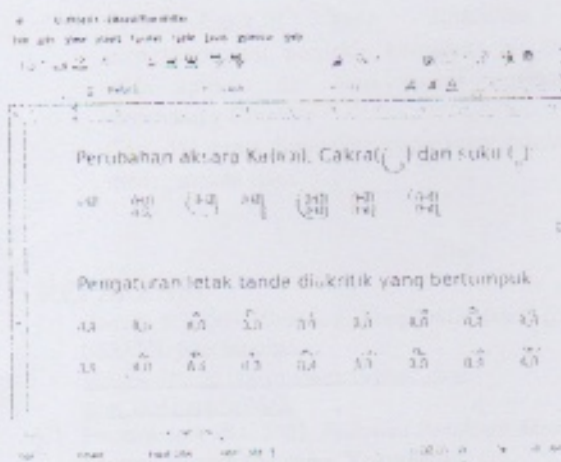
No	Input (Aksara Latin)	Hasil keluaran (Aksara Jawa)	Ket.
1	Konsonan	<i>Sigegan</i>	Sesuai
2	Konsonan + a	Konsonan	Sesuai
3	Konsonan+vokal selain a	Konsonan + <i>sandhangan swara</i>	Sesuai
4	Konsonan + r + a	Konsonan + <i>cakra</i>	Sesuai
5	Konsonan + r + vokal selain a	Konsonan + <i>cakra</i> + <i>sandhangan swara</i>	Sesuai
6	Konsonan+ y + a	Konsonan+ <i>pengkal</i>	Sesuai
7	Konsonan + y + vokal selain a	Konsonan+ <i>pengkal</i> + <i>sandhangan swara</i>	Sesuai
8	A	<i>Ha</i> / aksara swara a	Sesuai
9	Vokal selain a	<i>Ha</i> + <i>sandhangan swara</i> / aksara swara	Sesuai
10	Angka	<i>Angka</i>	Sesuai
11	.	<i>pada tingsu</i>	Sesuai
12	,	<i>pada lungsi</i>	Sesuai
13	: atau : atau ' atau "	<i>pada pangkot</i>	Sesuai
14	~ \	<i>pada adleg-adleg</i>	Sesuai
15	~ /	<i>pada windu</i>	Sesuai
16	- _	<i>pada andhap</i>	Sesuai
17	- -	<i>pada madya</i>	Sesuai
18	- +	<i>pada luhur</i>	Sesuai

### 3.3. Carik-Pad

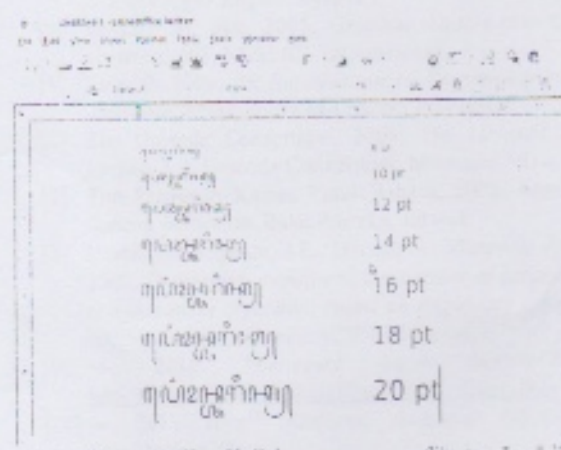
Carik-Pad adalah aplikasi transliterator yang dapat melakukan transliterasi teks beraksara Latin menjadi teks beraksara Jawa dan sebaliknya. Transliterasi ini ditujukan untuk digunakan oleh pengguna dengan pengetahuan penulisan aksara Jawa yang sangat terbatas. Transliterasi ini dibuat menggunakan bahasa pemrograman Java. Pola komponen suku kata bahasa Jawa dalam aksara Jawa dan aksara Latin diubah menjadi regex yang kemudian digunakan dalam proses penguraian pada saat transliterasi. Hasil penguraian kemudian dipetakan ke aksara tujuan berdasarkan tabel konversi.

Transliterasi ini mempunyai beberapa batasan. Pertama, transliterator tidak menggunakan aksara *murda* dan aksara *swara* pada proses transliterasi Latin-Jawa. Kedua, hasil transliterasi Jawa-Latin hanya berupa komponen penyusun suku kata bahasa Jawa. Hal ini dapat dimanfaatkan pengguna untuk membedakan aksara Jawa pada teks asal.

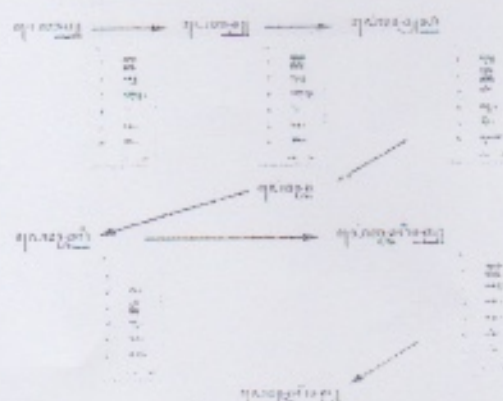




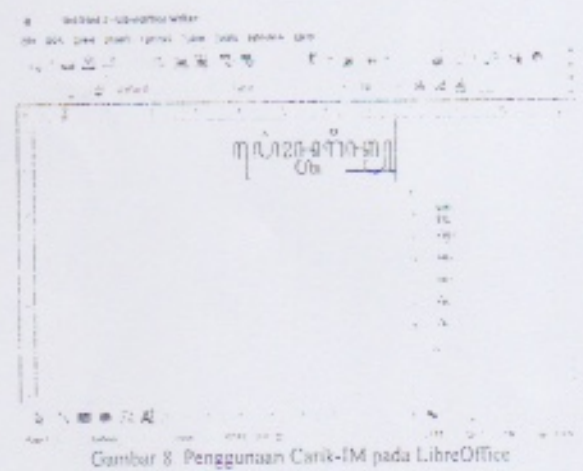
Gambar 5. Contoh perilaku kompleks aksara Jawa



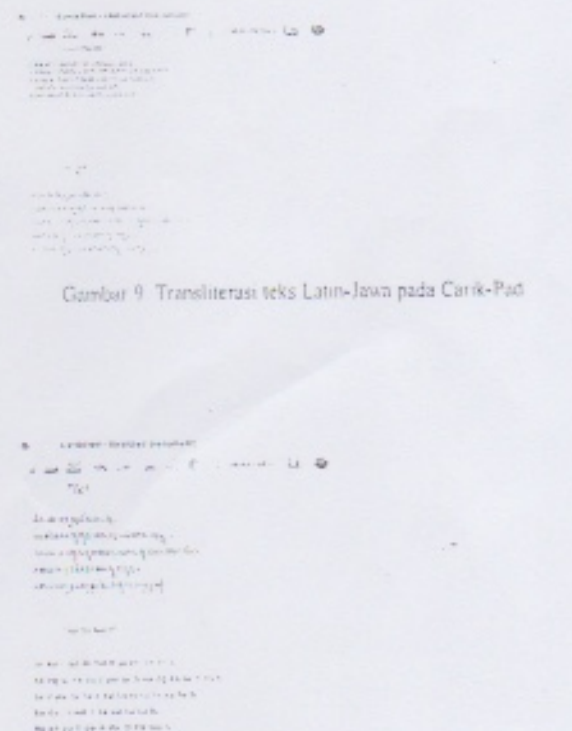
Gambar 6. Tampilan Carik ttf dalam beberapa ukuran



Gambar 7. Proses pemilisan "Font Carik" menggunakan Carik-IM



Gambar 8. Penggunaan Carik-IM pada LibreOffice



Gambar 9. Transliterasi teks Latin-Jawa pada Carik-IM

Gambar 10. Transliterasi teks Jawa-Latin pada Carik-IM

## 4. Kesimpulan

- Penelitian ini dapat menghasilkan sebuah font aksara Jawa Carik.ttf, dan dua jenis transliterator aksara Jawa. Transliterator yang dimaksud adalah Carik-IM input method yang dapat digunakan untuk transliterasi aksara Latin-Jawa dan Carik-Pad sebuah aplikasi yang dapat melakukan transliterasi aksara Latin-Jawa dan sebaliknya.



- b. Font Carik.ttf yang dihasilkan dapat mengakomodasi perilaku kompleks aksara Jawa pada aplikasi dan atau sistem operasi yang mendukung Graphite.
- c. Carik-IM dapat digunakan pada sistem operasi yang menggunakan IBus.

## REFERENSI

- [1] Correll, S., 2003, *Graphite: Frequently Asked Questions*, NSRI SIL International.  
[http://scripts.sil.org/cms/scripts/page.php?item\\_id=GraphiteFAQ](http://scripts.sil.org/cms/scripts/page.php?item_id=GraphiteFAQ).
- [2] Darusaputra, dkk., 2003, *Pedoman Penulisan Aksara Jawa*, Yayasan Pustaka Nusantara, Yogyakarta.
- [3] Habibi, M., 2004, *Java Regular Expressions: Taming the java.util.regex Engine*, Apress.
- [4] Hosken, M., dkk., 2007, *Graphite Description Language Version 2.003*, NSRI SIL International.
- [5] Linz, P., 2006, *An Introduction to Formal Language and Automata*, 4<sup>th</sup> ed, Jones and Bartlett Publishers.
- [6] The Unicode Consortium, 2009, *The Unicode Standard Version 5.2*, Unicode Consortium, Mountain View, CA.
- [7] Tim Penyusun Kamus Pusat Bahasa, 2008, *Kamus Besar Bahasa Indonesia*, Balai Pustaka, Jakarta.
- [8] Utami, E., Istiyanto, J.E., Hariati, S., Marsono, Ashari, A., 2009, *Developing transliteration pattern of Latin character text document algorithm based on linguistics knowledge of writing Javanese script*, ICICI-BME, 1-6.
- [9] ---, 2010, *Intelligent Input Bus*, Wikipedia.  
[http://en.wikipedia.org/wiki/Intelligent\\_Input\\_Bus](http://en.wikipedia.org/wiki/Intelligent_Input_Bus)
- [10] ---, 2011, *Java™ Platform, Standard Edition 6 API Specification*, Oracle.  
<http://download.oracle.com/javase/6/docs/api/>
- [11] ---, 2011, *Oxford Dictionaries Online*, Oxford University Press.  
<http://www.oxforddictionaries.com>